# COMP 555 – Bioalgorithms
# Course Syllabus
# Fall 2014

## Bulletin Description

Bioinformatics algorithms. Topics include DNA restriction mapping, finding regulatory motifs, genome rearrangements, sequence alignments, gene prediction, graph algorithms, DNA sequencing, protein sequencing, combinatorial pattern matching, approximate pattern matching, clustering and evolution, tree construction, Hidden Markov Models, randomized algorithms.

## General Course Info

| | |
|---|---|
| Term: | FALL 2014 |
| Department: | COMP |
| Course Number: | 555 |
| Section Number: | 001 |
| | |
| Time: | T R, 2:00 – 3:15 |
| Location: | FB 009 |
| Website: | `http://www.cs.unc.edu/~prins/Classes/555/` |

## Instructor Info

| | |
|---|---|
| Name: | Jan Prins |
| Office: | FB 334 |
| Email: | `prins@cs.unc.edu` |
| Phone: | 919-590-6213 |
| Web: | `http://www.cs.unc.edu/~prins/` |
| Office Hours: | TBA |

## Textbook

An Introduction to Bioinformatics Algorithms
by Neil C. Jones and Pavel A. Pevzner
MIT Press © 2004, ISBN: 0262101068.

## Course Description

Computational methods are fueling a revolution in the biological sciences. Computers are already nearly as indispensable as microscopes for analyzing and interpreting biological data. As a result, two new multidisciplinary fields, bioinformatics and computational biology, have emerged. This course will explore computational methods and algorithmic principles driving this revolution. It will cover basic topics in molecular biology, genetics, and proteomics. The course also addresses basic computational theory and algorithms including asymptotic notation, recursion, divide-and-conquer approaches, graph algorithms, dynamic programming, and greedy algorithms. These fundamental concepts from computer science will be taught within the context of motivating problems drawn from contemporary biology. Example biological topics

include sequence alignment, motif finding, gene rearrangement, DNA sequencing, protein peptide sequencing, phylogeny, and gene expression analysis.

This course is suitable for both computer science and biology students at both undergraduate and graduate levels. Students taking this course should have some programming experience in a modern programming language.

## Target Audience

This course is intended for advanced undergraduate computer science majors and computer science graduate students interested in exploring bioinformatics and computational biology. It is also well suited for graduate students from biomedical Departments or Curricula such as Biostatistics and the Curriculum in Bioinformatics and Computational Biology who would like to better understand the algorithms underlying the computational tools that drive modern research in these areas.

## Prerequisites

All students are expected to have taken introductory courses in data structures and discrete math equivalent to COMP 410 and MATH 381.

## Goals and Key Learning Objectives

Students completing this class should

- have an understanding of key algorithms used in bioinformatics and know their capabilities and limitations
- have knowledge of algorithmic strategies that can be employed to design new algorithms for bioinformatics applications
- be able to analyze the correctness and asymptotic time complexity of algorithms
- have a working knowledge of concepts and terminology in molecular biology and understand the key challenges
- be able to write bioinformatics programs using python and python bioinformatics libraries

## Course Requirements

Students must complete assigned reading from the textbook specified in each lecture. Students will be assigned problem sets, with associated programming assignments to be completed on their own computers using the python programming language. There will be five homework assignments, a midterm, and a final exam.

## Key Dates

| | |
|---|---|
| Midterm: | Thursday October 9, 2014 (in class) |
| Final Exam: | Saturday December 6, 2014, noon – 3pm |

## Grading Criteria

The final grade will be based on the follow weighting factors:

| | |
|---|---|
| 5 – Problem Sets/Programming Exercises | 50% (10% each) |
| 1 – Midterm Exam | 23% |
| 1 – Final Exam | 25% |
| Class participation | 2% |

## Course Policies

- Attendance is expected, but no roll will be taken
- Problem sets will only be accepted at the beginning of class meetings on the date they are due.

The course final is given in compliance with UNC final exam regulations and according to the UNC Final Exam calendar.

## Honor Code

The UNC honor code is in effect for this course, which means that all submitted work must be your own. For the written assignments only, collaboration on strategy and approach to a problem is permitted and encouraged. However, you must write your own detailed solutions which may take into account your understanding of collaborative discussions. Good scholarship requires that all collaboration to be acknowledged. Thus, if you collaborate on the solution strategy of a problem set, I expect that you list your collaborators at the top of the first page.

Collaboration on tests (midterms, final) is a violation of the Honor Code. This includes discussion of questions on a midterm, or final with students that have not yet taken the test.

Using any unauthorized information sources on an exam is a violation of the honor code.

## Course Schedule

A course schedule and handouts from each lecture will be posted on the course website.

## Disclaimer

“The professor reserves to right to make changes to the syllabus, including project due dates and test dates. These changes will be announced as early as possible.”